Executive Summary

My capstone will be focusing on historic tornado and traffic data for Nashville. I wanted to see where peak hours of traffic were in comparison to when tornadoes typically strike, hopefully giving us insight into where there might be bottlenecks. The challenge comes from the fact that tornadoes move constantly. Assuming the data makes a distinction of a start and end point for the tornado, along with time of day for traffic data, getting all the data in a row could prove difficult.

Motivation

I came to this idea as a friend I knew lost his life in the 2020 Nashville tornado, and living in Tennessee, it was always a concern every year if our house was going to be destroyed. In the dataset I was able to find a tornado that devastated the university my brother and sister went to, which gave a tangibleness to the data.

Data Question

What times to tornadoes touch down in Nashville and where? What times are busiest for traffic that might make it hard if a tornado touches down?
I found a lot of awesome resources for specifically tornadoes and their effect on the U.S., but I was not able to find other people asking this specific question.

Research:
https://developers.arcgis.com/python/sample-notebooks/analyze-us-tornadoes/
https://www.spc.noaa.gov/wcm/#data

https://www.weather.gov/

Minimum Viable Product (MVP)

My MVP will consist of a presentation showing peak times of day for traffic and what counties are busiest, along with tornadoes and where they struck down, their size, along with their trajectory year to year. This should give a good picture of possible traffic issues that might pop up for future tornadoes. The intended audience is the Tennessee Government.

Schedule (through 4/30/2021)

1. Get the Data (4/16/2021)
2. Clean & Explore the Data (4/21/2021)
3. Create Presentation of my Analysis (4/23/2021)
4. Internal demos (4/26/2021)
5. Demo Day!! (4/30/2021)

Data Sources

Data:

https://hifld-geoplatform.opendata.arcgis.com/datasets/historical-tornado-tracks?geometry=-89.018%2C35.804%2C-84.349%2C36.579&orderBy=st&page=12&where=st%20%3D%20%27TN%27

https://www.arcgis.com/apps/webappviewer/index.html?id=075987cdae37474b88fa400d65681354

https://www.census.gov/data/tables/time-series/demo/popest/2010s-counties-total.html#par_textimage

Known Issues and Challenges

Explain any anticipated challenges with your project, and your plan for managing them. Be sure to include:
- *I'm working with three to four datasets from different sources, and they may not play nice with each other*

- *The datasets are shape files and csv files, so I will have to convert them to geodataframes*

- *County columns for the datasets are not consistent for the datasets, so making sure that all counties are covered is important*

- *Working with powerbi, there may need to be some backtracking to make sure everything in the cleaned dataset plays nicely when making the presentation*